Effects of Plurilingual Teaching on Grammatical Development in Early Foreign-Language Learning

HOLGER HOPP[1] AND DIETER THOMA[2]
[1] *Technische Universität Braunschweig, English Linguistics, Bienroder Weg 80, Braunschweig, 38106, Germany Email: h.hopp@tu-braunschweig.de*
[2] *Universität Mannheim, English Linguistics, Schloss EW266, Mannheim, 68131, Germany Email: thoma@uni-mannheim.de*

This article reports two intervention studies testing the effects of plurilingual teaching on grammatical development among primary-school students learning English as a foreign language (FL). In a pre–posttest control-group design, more than 200 9–10-year old majority language German and minority language students received plurilingual FL teaching (intervention group) or regular FL-only teaching (control group). Study 1 on the acquisition of *wh*-questions showed that systematic cross-linguistic comparisons of the FL with the majority language and minority languages facilitate acquisition of object questions. In Study 2 on passives, the intervention and the control groups both demonstrated comparable gains. We suggest that plurilingual teaching has advantages when the majority language differs from the target language (Study 1) yet not when a phenomenon is comparable across languages (Study 2). In neither study did learners show generalization to related grammatical phenomena. Finally, majority language and minority language students did not perform differently, which suggests that plurilingual FL teaching is suitable for all FL learners. These findings demonstrate that plurilingual FL teaching facilitates grammatical development by increasing learners' awareness of cross-linguistic similarities and differences.

Keywords: plurilingual teaching; pedagogical translanguaging; foreign language teaching; transfer; grammar; achievement

ACROSS MOST COUNTRIES, CLASSROOMS are becoming more heterogeneous in terms of student backgrounds, qualifications, and needs. In the context of foreign language (FL) instruction, an increasing number of students have previously acquired more than one language, and they thus potentially differ in their starting points and the trajectories of their FL development from monolingual students.

Plurilingual approaches to FL teaching capitalize on such linguistic diversity and advocate integrating previously acquired languages in FL instruction with the aims of developing plurilingual competence that spans learners' entire linguistic repertoires (e.g., Krumm & Reich, 2013), or facilitating FL learning by virtue of positive transfer, cross-linguistic inferencing, and increased metalinguistic awareness (e.g., Hufeisen & Neuner, 2003; for a review, see Cenoz & Gorter, 2013). To date, plurilingual approaches to FL teaching have predominantly been applied to and assessed for early balanced bilinguals living in diglossic societies, such as Catalonia and Canada (for a review, see Cenoz, 2003), or in further FL learning where learners can build on their experiences in learning the first FL in instructed

The Modern Language Journal, 0, 0, (2021)
DOI: 10.1111/modl.12709
0026-7902/21/1–20 $1.50/0

settings (for a review, see Fäcke & Meissner, 2019). Across many Western societies, a large number of children grow up speaking a heritage or minority language on top of the societal majority language (Montrul, 2016). Typically, they have varying skills in the minority languages, and in each classroom, a large variety of minority languages are represented. Much less is known about the effects of plurilingual teaching on development in the first FL among such minority language students with a migration background. The present contribution probes effects of plurilingual FL teaching in the context of early FL learning in mixed primary-school classrooms consisting of monolingual majority language speakers as well as bilingual minority language speakers.[1]

In these contexts, plurilingual FL teaching can act as a potential resource to facilitate FL learning in at least two ways. First, plurilingual FL teaching can address the specific (meta)linguistic resources of minority language students and, thus, partially compensate for the lower FL achievement often reported for this group (e.g., Maluch, Neumann, & Kempert, 2016). Second, plurilingual FL teaching can potentially boost FL development for both majority language and minority language students, since it addresses their resources as "emergent multilinguals" (Turnbull, 2018). Both compensatory and additive effects of plurilingual FL teaching inform curricular guidance to include previously acquired languages in (early) FL instruction in many countries and states (e.g., Niedersächsisches Kultusministerium, 2018).

Some initial studies in these mixed classroom settings have reported positive effects of plurilingual FL teaching on vocabulary acquisition (Busse et al., 2020). In order to go beyond item-specific learning of vocabulary, the present study examines effects of plurilingual FL teaching on grammatical development. In addition, the present study focuses on early FL learning in view of findings that effects of minority languages on FL learning decrease as students get older and transition into secondary schooling (Hopp, Vogelbacher et al., 2019; Maluch et al., 2016).

We report two studies conducted in a pre–posttest control group design with 9–10-year-old 4th-grade students at regular German primary schools. The intervention group was taught according to the principles of pedagogical translanguaging (PTL; Cenoz & Gorter, 2017, 2020)—a specific implementation of plurilingual FL teaching—while the control group took part in regular English lessons conducted in the target language only. The first study targets *wh*-questions in English. English *wh*-questions differ in word order from their counterparts in the majority language, German, yet share characteristics with many languages spoken as minority languages among the students. The second study deals with passives that share many characteristics with the majority language as well as with many minority languages. Comparing *wh*-questions and passives allows us to investigate whether effects of plurilingual FL teaching, if any, are modulated according to the differences and similarities between the grammars of the languages involved. In each study, we assess effects of PTL on the target structure as well as a syntactically related structure that was not subject to instruction, in order to examine whether learning effects generalize across syntactic constructions. Further, we test whether majority language learners and minority language learners demonstrate differential learning gains to address whether plurilingual FL teaching has compensatory or additive effects in grammatical learning.

PLURILINGUAL FOREIGN LANGUAGE TEACHING AND PEDAGOGICAL TRANSLANGUAGING

FL teaching typically centers around communicative activities conducted in the target language only so as to best emulate the conditions of immersive naturalistic language acquisition leading to native-like attainment. However, recent proposals have shifted the perspective away from viewing FL learning as aiming to approximate monolingual native speaker proficiency.

Based on psycholinguistic and sociolinguistic work illustrating the interconnectedness of linguistic knowledge and resources in bilinguals and multilinguals, Cook (2013) coined the term "multicompetence" to describe the resources of multilingual speakers that encompass competences across languages and render them fundamentally different from monolinguals (see also Grosjean, 1989). In a broader perspective, the notion of 'translanguaging' addresses the entire linguistic repertoire of bilingual and multilingual speakers and softens the perception of boundaries between languages in multilingual discursive practices (García & Li Wei, 2014).

Against this backdrop, plurilingual approaches to FL teaching formulate methods and practices that include references to other previously learned languages (Candelier & Kervran, 2018; Martinez, 2015). One such approach, PTL (Cenoz & Gorter, 2017, 2020), centers on teacher-designed instructional strategies that include the

use of more than one language in the classroom in input and output activities, implicit and explicit language comparisons, the use of cognates, or multilingual discursive tasks (see also Cenoz, 2019).

Plurilingual teaching approaches derive their evidentiary base from studies on FL achievement that compare previously monolingual and bilingual learners either conducted in laboratory settings or in classroom contexts (for a review, see Hirosh & Degani, 2018). Both lines of research often report advantages of early balanced bilingual learners over monolingual learners (for a review, see Cenoz, 2013). Similarly, studies in instructed language acquisition yield advantages for students who previously learned another FL at school over those without previous language learning experiences (e.g., Haenni Hoti et al., 2011; though see Berthele & Udry, 2019).

It is less clear whether multilingual advantages can also be observed in instructed FL learning of minority language learners in majority-language contexts. Many studies on early FL learning conducted in European countries do not report any differences between minority language and majority language students in standardized comprehension or writing tests of the FL (e.g., Schoonen et al., 2003; Wilden & Porsch, 2016), and some studies even report significant disadvantages for minority language students as a group already at the primary school level (e.g., Elsner, 2007). However, when individual differences in cognitive skills and social background are controlled for, some selective advantages of minority language learners emerge (e.g., Hesse, Göbel, & Hartig, 2008; Maluch et al., 2015), yet multilingual boosts do not extend across all groups of learners (Göbel & Vieluf, 2017). In a recent study on German 9–10-year-old 3rd- and 4th-grade primary school minority and majority language students, Hopp, Vogelbacher et al. (2019) reported that both cognitive and linguistic skills in the minority language contribute significantly to vocabulary and grammar skills in FL English. When all cognitive and linguistic factors were statistically controlled for, minority language speakers had advantages over majority language speakers in the FL, in particular, if they had a large vocabulary in the minority language (see also Maluch & Kempert, 2017; Fleckenstein, Möller, & Baumert, 2018). These findings suggest that having acquired a minority language may confer selective benefits in FL learning (see also Bonifacci et al., 2017).

Part of the reason for the mixed findings for minority language students in immigration contexts may stem from the lack of systematic pedagogical approaches and teaching methods that refer to previously acquired language among FL learners and could thus stimulate their linguistic resources in fostering FL development. As a case in point, Hopp, Vogelbacher, et al. (2019a) reported that beneficial effects of previously acquired languages on FL skills wane from 3rd to 4th grade in primary school (see also Maluch et al., 2016). This may be taken to indicate that traditional (i.e., target-language only) FL teaching does not sufficiently address the linguistic resources of minority language learners in FL learning.

Emerging research on the effects of plurilingual teaching in both bilingual and immigration contexts suggests that the inclusion of other languages in FL teaching may lead to learning advantages over FL-only instruction. In French immersion programs in Canada, 7–8-year-old 2nd grade students who were taught about affixation in French and English using both languages outperformed a control group that received instruction only in one language (Lyster, Quiroga, & Ballinger, 2013). For morphological awareness, Basque–Spanish 5th- and 6th-grade bilingual learners of English (ages 10–11) who were taught about derivation and compounding using plurilingual examples across all three languages demonstrated more consistent gains in morphological awareness tasks after a 12-week intervention than a control group who received instruction only in each respective language (Leonet, Cenoz, & Gorter, 2020). For vocabulary learning, a plurilingual intervention study with 11–14-year-old Spanish- and English-speaking students in grades 6–8 in the United States focused on the use of English–Spanish cognates to foster vocabulary and reading comprehension (Arteagoitia & Howard, 2015). Among the students, knowledge of cognate words in Spanish positively correlated with English vocabulary and reading comprehension, which suggests that the students accessed their knowledge of Spanish in FL tasks.

For mixed primary-school classrooms comprising majority language and minority language learners, Busse et al. (2020) reported positive effects of plurilingual teaching on vocabulary acquisition in an instructional treatment spanning five 45-minute lessons among 42 German 3rd-grade minority and majority language students aged 8–10 years. Half of the students participated in several activities involving the use of minority language words in the intervention group. The other half of students in the control group received regular English-only teaching. In productive and receptive assessments on the novel vocabulary items at the end of each lesson, the

intervention group consistently outperformed the control group. In this study, the teaching of the intervention and the control group differed along many dimensions—for example, by also including motivating elements to foster positive affect in the intervention group—which leaves open whether the learning advantage of the intervention group is owing to the inclusion of other languages or other elements of the treatment. Further, the study did not consider differences in learning gains between majority language and minority language students, so that the group differences cannot be related to the activation of the linguistic repertoires of the students.

In addition, research for minority language students so far is limited to direct effects of interventions involving plurilingual teaching on the acquisition of word-level characteristics or individual vocabulary items. We do not know if the scope of effects extends beyond vocabulary to more abstract, rule-governed aspects of language in grammar.

Research on instructed second language (L2) acquisition in adults suggests that reference to first language (L1) grammar can promote L2 or FL grammatical development. In a series of studies, McManus and Marsden (2017, 2019a, 2019b) assessed the effects of explicit metalinguistic information about differences between the L1 and the FL grammar on the acquisition of tense and aspect distinctions in French. In interventions lasting 3–4 weeks for 45 minutes each week, the studies compared groups of university-level FL learners receiving no explicit instruction on aspectual distinctions with groups being instructed on L2 aspectual distinctions only and groups being taught about contrasts in aspectual distinctions in the L1 and the L2. On top of explicit information, the groups practiced the relevant aspectual distinctions in various activities in the L1 and the L2. Across studies, the groups that received additional contrastive information about the realization of form-to-function mappings in the L1 outperformed the other groups in immediate and delayed posttests involving comprehension or oral production. Hence, instructional treatments that increased learners' awareness of similarities and differences in L1 and L2 form–function mappings led to greater learning gains in instructed FL learning in attenuating negative influence from the L1. Although these studies were not carried out in the context of plurilingual FL teaching, their findings suggest that reference to other languages than the FL may positively affect grammatical development in the FL. Of course, the methods of explicit instruction of L1 and L2 grammatical rules do not easily translate to early FL learning among low-proficiency young learners who have limited metagrammatical knowledge and terminology at their disposal.

In sum, emerging research reports positive effects of plurilingual approaches to FL learning; yet it is an open question whether and how plurilingual teaching affects grammatical development in early FL learning, and whether it has differential impacts among minority language and majority language learners in mixed classrooms.

RESEARCH QUESTIONS AND HYPOTHESES

Against the backdrop of previous research, we carried out two experiments on the effects of PTL in early FL learning at primary school. Previous studies capitalized on cross-linguistic similarities at the word level, such as the use of cognate words and internationalisms (Arteagoitia & Howard, 2015; Busse et al., 2020) or analogous word-formation processes in affixation or compounding (Leonet et al., 2020; Lyster et al., 2013). In contrast, research on instructed L2 acquisition emphasizes the usefulness of explicit instruction about the L1 for cross-linguistic differences (McManus & Marsden, 2019a, 2019b). With a view to effects of PTL on grammatical development, the present study thus includes two experiments, the first focusing on grammatical differences between the FL and the majority language, and the second dealing with a phenomenon for which the FL and the majority language are similar. We pose the following three research questions.

RQ1. Does PTL by virtue of the use of contrastive language activities promote the acquisition of foreign-language grammar?

We assess effects of PTL in a control-group design in which the instructional treatments of the PTL group and the control group are maximally similar and differ only by virtue of the addition of PTL elements. During the lessons, these PTL elements take up time that is devoted to further practice of the FL in the control group. Hence, we test whether PTL overcompensates for less time on task in FL grammatical development.

RQ2: Do effects of PTL extend across different grammatical phenomena?

We test the scope of PTL in two different ways. First, we conduct two studies to assess whether

PTL shows differential effects for phenomena that are different or similar between the FL and the majority language—namely, *wh*-questions and passives. Second, within each study, we probe whether potential effects of PTL on the phenomenon of instruction transfer to syntactically related constructions that were not taught in the instructional treatment.

RQ3: Are there differences in the degrees to which PTL affects the grammatical development of majority language and minority language learners?

In each study, we compare learning gains of majority language and minority language students. If PTL primarily derives its efficacy from activating the linguistic resources of learners across their entire linguistic repertoire, we expect minority language students to benefit to a larger degree from PTL than majority language students. Since the particular PTL activities include the use of the majority language (German), we can also predict that majority language students will potentially benefit from PTL.

The studies test the effects of focused interventions spanning two 45-minute lessons each. Critically, the specific interventions presented here were embedded in a 6-month intervention project on plurilingual early FL teaching, such that students in the PTL intervention group did not perceive the treatments as unusual, novel, or particularly salient. This way, the present experiments can isolate the effects of PTL versus other potential differences implicated in plurilingual FL teaching compared to regular FL instruction.

THE STUDIES

Participants

Overall, the participants comprised 258 4th-grade students at four different German public primary schools in medium-sized cities in the state of Lower Saxony. At each school, two intact classes in 4th grade took part in the study in the academic years 2018–2019 (Cohort 1) and 2019–2020 (Cohort 2). Each class comprised at least 33% of students who spoke minority languages on top of the majority language. One class at each school was randomly assigned to PTL group, and one class was assigned to the control group. All students had been learning English in two 45-minute lessons per week since Grade 3, that is, for about 1 year and 3 months before the start of data collection.[2] Table 1 gives background data about the participants in the two cohorts, split by group as well as by majority language and minority language learners. On top of having German as their dominant language, the minority language students, most of whom were born in Germany, had proficiency in one of the following 15 languages spoken in their families or homes: Albanian, Arabic, Chinese, Croatian, Farsi, Greek, Indonesian, Italian, Kurdish, Polish, Romanian, Russian, Spanish, Turkish, and Vietnamese. Our sample thus comprised a varied set of minority languages.

Several weeks before the intervention, all students completed tasks on cognitive processing and language proficiency. For cognitive processing, we administered a nonverbal IQ test (CFT 20-R; Weiβ, 2006) and assessed working memory using a forward digit span task (adapted from HAWIK-IV; Petermann & Petermann, 2008). For language proficiency in German, students completed a standardized 32-item picture naming task (Cross-Linguistic Lexical Task, CLT; Rinker & Gagarina, 2014), and, for English, they completed a receptive vocabulary task (British Picture and Vocabulary Scale, BPVS3; Dunn, Dunn, & Styles, 2009) and a receptive grammar task (Test for Reception of Grammar, TROG2; Bishop, 2003). As shown by the similar means in Table 1 across all cognitive and linguistic tasks, all groups were comparable. In fact, one-way ANOVAs with the respective test scores as dependent variables and either group or speaker language status as independent variables did not return any significant differences in any comparison for Cohort 1 (all ps $> .11$), except for lower German proficiency for the minority language students, $F(1,121) = 19.027$, $p < .001$. In Cohort 2, the PTL group and the control group did not differ on any variable (ps $> .38$), yet the majority language speakers were younger than the minority language speakers, $F(1,136) = 7.020$, $p = .009$; had higher IQ scores, $F(1,136) = 7.510$, $p = .007$; and had higher proficiency in German, $F(1,136) = 10.589$, $p < .001$. On all other cognitive and linguistic variables, they were matched (ps $> .42$). We also assessed proficiency in the minority language using the CLT among the minority language students. The scores for the production task in the minority language showed that all students achieved considerably lower scores in their minority language ($M = 11.8$; $SD = 7.9$) than in German (see Table 1), indicating that they were dominant in the majority language.

Study Design and Intervention Procedure

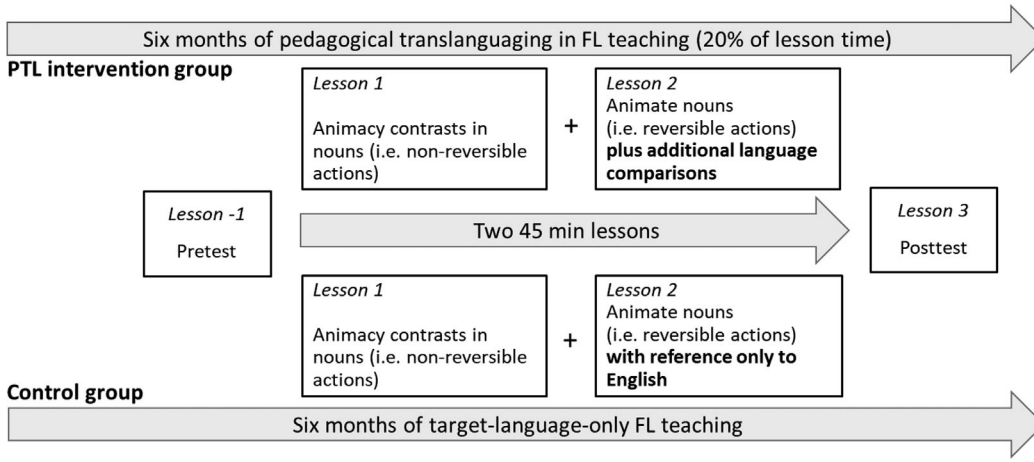
The study was embedded in a larger research project on effects of plurilingual FL teaching on

TABLE 1
Participant Characteristics

	Cohort 1 (Study 1)				Cohort 2 (Study 2)			
	PTL Group		Control Group		PTL Group		Control Group	
	Majority Language	Minority Language	Majority Language	Minority Language	Majority Language	Minority Language	Majority Language	Minority Language
n	39	36	28	27	37	29	37	25
Age in years	9.6 (0.5)	9.9 (1.0)	9.6 (0.5)	9.6 (0.6)	9.5 (0.4)	9.7 (0.6)	9.6 (0.4)	9.7 (0.6)
Length of exposure to German in years		8.6 (4.2)		8.0 (4.3)		7.0 (3.0)		8.1 (2.9)
Nonverbal IQ (CFT-20R)	104.5 (16.4)	99.1 (24.7)	100.0 (15.8)	97.0 (15.3)	107.0 (12.3)	101.2 (14.7)	108.3 (14.4)	101.0 (14.8)
Working memory (digit span)	7.0 (1.3)	7.1 (1.4)	6.8 (1.7)	6.8 (1.1)	7.0 (1.3)	7.3 (1.2)	7.5 (1.8)	6.8 (1.4)
German proficiency (CLT production)	30.3 (1.3)	29.4 (1.9)	30.8 (1.0)	29.3 (1.7)	30.6 (1.1)	28.2 (3.6)	30.0 (1.5)	28.4 (3.6)
English receptive vocabulary (BPVS3)	43.6 (6.7)	41.4 (8.1)	40.1 (5.6)	40.7 (7.3)	44.7 (5.7)	45.0 (8.0)	44.1 (6.1)	43.4 (7.2)
English receptive grammar (TROG2)	44.9 (9.6)	41.9 (10.2)	40.9 (7.7)	42.5 (9.2)	43.9 (8.0)	47.4 (10.8)	46.1 (10.5)	41.4 (10.4)

Note. PTL = pedagogical translanguaging; CFT-20R = Culture Fair Intelligence Test; CLT = cross-linguistic lexical task; BPVS3 = British Picture and Vocabulary Scale; TROG2 = Test for Reception of Grammar. Values reported are means (standard deviations in parentheses).

FIGURE 1
Data Collection Procedure and PTL versus Control Teaching Activities



Note. PTL = pedagogical translanguaging; FL = foreign language.

the development of general (meta-)linguistic and cognitive skills (see Hopp et al., 2020). Figure 1 schematically illustrates the procedure. As part of the project, the classes in each cohort received English lessons with PTL elements (PTL group) or regular English lessons (control group) for 6 months, from after the autumn break to before the Easter break. The PTL lessons aimed to engender the appreciation of linguistic diversity, provide opportunities for the use of other languages in FL learning, and promote the discovery of language similarities and contrasts as well as foster metalinguistic awareness (for details, see Hopp et al., 2020). In these English lessons, approximately 20% of all activities comprised PTL activities including multilingual songs; the use of words from different languages in vocabulary exercises; as well as linguistic comparisons of English vocabulary or spelling with the majority language, German, and the minority languages spoken by students in each class. In addition, the group regularly engaged in activities fostering phonological awareness using nonwords or English words. All PTL activities blended with the textbook used in class (*Playway 4*; Gerngross, Puchta, & Becker, 2013). Importantly, the students in the PTL group were thus already familiar with PTL, so that the intervention was not unusual or novel for them in terms of references to languages other than English or the contrastive and comparative nature of the tasks. The control group received the same teaching, yet without the PTL components. Instead, the students in the control group completed longer or additional activities from the textbook or engaged in similar activities to the intervention group that, however, were conducted in the target language only. At each school, both the PTL and the control group were taught by the same teacher, namely, a project member who was a qualified English teacher. The regular English teachers were present during the English lessons and supervised the teaching.

The instructional treatments in Studies 1 and 2 were embedded in the textbook unit on birthdays. In the treatments, the vocabulary of the respective textbook unit was used, so that the treatments blended with the regular teaching. Each treatment comprised two 45-minute lessons and was preceded by a pretest in the lesson before the first treatment lesson and a posttest in the subsequent English lesson. Treatment fidelity was ensured by fixed teaching scripts and joint preparation of the materials and lessons by the teachers. Lesson documentations and debriefings showed that the teachers implemented the lessons in highly comparable ways across classes and schools.

The study was authorized by the state's school board, and parents gave written informed consent for their children to participate in the tests. All children were informed about the tests and gave oral consent prior to each testing session.

Study 1: Wh-Questions and Relative Clauses

Target Phenomenon. *Wh*-questions and relative clauses present challenges for language learners since their surface word order does not transparently reflect canonical syntactic or thematic roles—the first noun phrase is not necessarily the subject or agent in the clause. In object

questions like *Which boy does the man see?* the first noun phrase is the object of the verb. As a consequence, both monolingual and bilingual child L1 learners, as well as early and late L2 learners, experience delays in the acquisition of object questions and object relative clauses and initially overinterpret them as subject-first strings. Such difficulty has been argued to follow from a canonical mapping strategy between syntactic function and syntactic position of arguments (Canonicity Hypothesis; Friedmann & Novogrodsky, 2004). Even among adult monolinguals, object questions and object relative clauses require additional processing effort in that an initial subject-first preference needs to be revised to an object order (e.g., DeVincenzi, 1991). For child FL learners of English, Hopp, Steinlen et al. (2019) found that 4th-grade 9–10-year-old students in regular German public primary schools had low comprehension accuracy of object questions and relative clauses, interpreting them as subject-first orders around 70% of the time (see also Roesch & Chondrogianni, 2016, for child L2 data).

Part of the difficulty of L1 German learners may result from differences in how English and German realize *wh*-questions and relative clauses. In English, differences between subject and object strings are overtly signaled in word order differences (1).

(1) a. *Which animal*$_i$ t_i *bites the horse?* (subject question)
b. *Which animal*$_i$ *does the horse bite* t_i*?* (object question)
c. *The animal that*$_i$ t_i *bites the horse.* (subject relative clause)
d. *The animal that*$_i$ *the horse bites* t_i. (object relative clause)

In formal linguistic analyses, these word order differences derive from movement of a *wh*-phrase or a relative pronoun to clause-initial position from its thematic base position (e.g., Chomsky, 1981), indicated by coindexed traces in (1).

In German, subject and object *wh*-questions have identical surface word order, as seen in the German versions of the English sentences in (1) in (2). In main clause *wh*-questions, the verb-second (V2) property of German allows only one preverbal constituent, which leads to identical word orders for subject and object questions. In relative clauses, the underlying object–verb (OV) word order of German embedded clauses creates a similar ambiguity in that the word order of subject and object relative clauses is identical. Instead of word order, German uses case marking on prenominal determiners to mark syntactic function. However, many determiner forms are ambiguous: For example, the article for the singular neuter noun *Pferd* 'horse' is the same in nominative and accusative in (2).

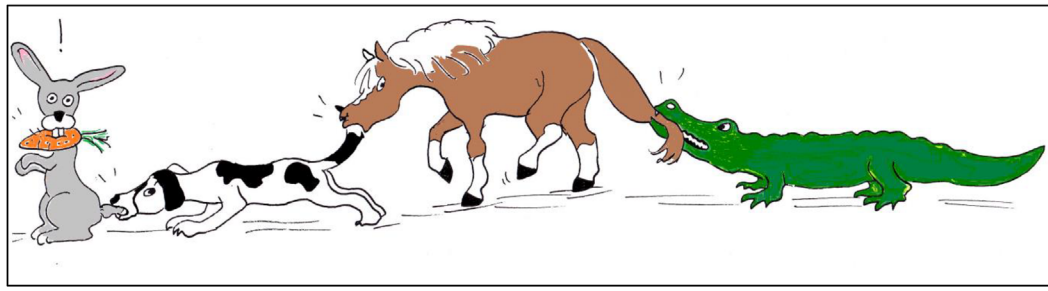
(2) a. *Welches Tier*$_i$ *beißt*$_j$ t_i *das Pferd* t_j *?* (subject question)
b. *Welches Tier*$_i$ *beißt*$_j$ *das Pferd* t_i t_j*?* (object question)
which animal bites the horse
c. *Das Tier, das*$_i$ t_j *das Pferd beißt.* (subject relative clause)
d. *Das Tier, das*$_i$ *das Pferd* t_i *beißt.* (object relative clause)
the animal that the horse bites

Complete surface word order overlap between subject and object *wh*-questions and relative clauses is specific to languages with V2 or OV grammars such as Afrikaans, Dutch, or German. All other languages, including the minority languages spoken by the students in the study, make word order distinctions between subject and object questions and/or relative clauses.

In consequence, we employed these contrasts between English and the majority language German as well as the similarities between English and minority languages to test whether PTL has beneficial effects on how students comprehend *wh*-questions, which were subject to instruction in the treatment, as well as relative clauses, which were not subject to instruction.

Instructional Treatments. Both the PTL group and the control group received largely identical treatments, spanning two 45-minute English lessons. In the first lesson, *wh*-questions were introduced by the teacher in the context of a birthday party, in which people give and receive presents. Using picture cards, the teacher exemplified the contrast between subject and object *wh*-questions by relying on animacy as a cue to distinguish between the two (*Who brings the cake?* vs. *What does John bring?*). Students practiced both types of questions in comprehension (picture identification) and production tasks. In the second lesson, the teacher created the context of an animal birthday party, in which animals give each other toy animals as presents. In this context, subject and object questions differed only in word order (*Who brings the cat?* vs. *Who does the cat bring?*). Using similar activities as in Lesson 1, the students practiced the comprehension and production of questions. Subsequently, the teacher in the PTL group wrote down subject and object questions on the blackboard in English, and she illustrated questions in German. She asked minority language students to give examples of questions from other languages spoken by them if they wanted, noting after each example that the word order differs between subject and object

FIGURE 2
Display for Sentences Shown in (3) [Color figure can be viewed at wileyonlinelibrary.com]



questions. This way, the teacher illustrated that subject and object questions differ in word order in English and other languages, while they share the same word order in German. Toward the end of the lesson, students filled in a worksheet by writing down questions in English, German, and other languages. In the control group, the teacher only wrote down English questions on the blackboard, and the students practiced the use of English questions further. They completed the same worksheet by filling in several subject and object questions in English. In both groups, the teaching aimed to illustrate the word order differences between subject and object questions, yet the treatment in the PTL group included reference to the majority language German and the minority languages spoken in class.

Testing. The testing materials were adapted from Hopp, Steinlen et al. (2019) and based on Rankin (2014). The test comprised 10 quadruplets of items with familiar animal names and the verbs *bite* and *catch* as in (3).

(3) a. *Which animal bites the horse?*
b. *Which animal does the horse bite?*
c. *The animal that bites the horse.*
d. *The animal that the horse bites.*

As in (3), the questions used complex *wh*-phrases (*which animal*) and relative clauses with *the animal* as the head noun. The items were distributed across 10 displays as in Figure 2, in which animals performed identical actions on one another.

For each display, there was one *wh*-question, one relative clause, and one filler question, targeting the action or shape of an animal depicted (e.g., "Which animal has a carrot?"). In all, the test encompassed 10 questions (5 subject and object questions each), 10 relative clauses (5 subject and object relative clauses each), as well as 10 filler items.

Testing Procedure. We created two lists of each test, counterbalancing the order of lists between pretest and posttests across intervention and control group. The tests were completed in class, with the teacher leading the task. Each student received a booklet with the pictures for all items. The teacher explained the task in German and answered questions. She told the students to circle the target animal that would be the answer to the question. The test began with one practice sentence and one filler sentence. For each display, the teacher first named all animals from left to right and then read the experimental sentence twice at a slow pace. Then students circled the answers. There was no feedback, and the regular class teacher walked around in the classroom to make sure students did not miss items or copy from each other. In all, the main test took approximately 25 minutes.

Participants. One hundred twenty-five students from Cohort 1 participated in the treatment on *wh*-questions. Data from both the pretest and the posttest were obtained from 115 students, 62 students in the PTL group (33 minority language learners), and 53 students in the control group (26 minority language learners).

Analysis and Results. We excluded one student from the control group, since the student was bilingual with English as one of the first languages. For the remaining 114 participants, we coded their responses for accuracy and type of mistakes. All data were analyzed using mixed logistic regression modelling (see also Kasprowicz, Marsden, & Sephton, 2019), using the glmer function from the lme4 package, version 1.1-21 (Bates et al., 2019) in R Studio (R Development Core Team, 2020). As fixed effects, we entered test (pretest vs. posttest),

TABLE 2
Study 1: Mean Accuracy Scores and Choices in % (Standard Deviation in Parentheses) by Test, Construction, and Group ($n = 114$)

		Pretest			Posttest		
Type	Order	Accurate	Object/Subject	Named	Accurate	Object/Subject	Named
PTL group ($n = 62$)							
Fillers		89.8 (10.8)			94.4 (9.5)		
Questions	Subject	87.1 (17.4)	11.9 (16.8)	1.0 (4.3)	88.7 (18.3)	11.0 (18.4)	0
	Object	16.8 (22.4)	81.6 (22.7)	1.3 (6.1)	31.6 (34.1)	66.8 (34.9)	1.6 (6.6)
Relative clauses	Subject	91.9 (15.1)	6.5 (14.4)	1.3 (5.0)	85.5 (20.2)	12.6 (19.6)	1.6 (5.5)
	Object	26.8 (22.6)	66.8 (24.3)	5.5 (9.7)	26.5 (28.7)	72.3 (29.9)	1.0 (4.3)
Control group ($n = 52$)							
Fillers		85.6 (13.5)			91.0 (13.0)		
Questions	Subject	86.5 (20.9)	10.4 (17.9)	3.1 (9.2)	89.6 (17.0)	9.2 (17.0)	1.1 (4.7)
	Object	16.5 (19.7)	80 (21.3)	2.3 (7.6)	22.3 (30.1)	75.8 (31.0)	1.5 (6.7)
Relative clauses	Subject	87.3 (20.6)	9.2 (14.0)	2.7 (8.9)	83.8 (23.8)	14.2 (21.8)	1.2 (4.7)
	Object	26.2 (25.8)	67.3 (30.4)	4.6 (10.9)	26.9 (31.4)	70.0 (33.3)	2.7 (6.9)

Note. PTL = pedagogical translanguaging; Object/Subject = the respective other interpretation of the string (e.g., subject interpretation of object question); Named = animal named in question.

order (subject vs. object), group (PTL vs. control), and language status (majority language vs. minority language). The contrasts were treatment coded, with pretest for test, subject for order, and control group for group as reference levels. Models were kept maximal in terms of the random-effect structure. When this maximal model did not converge, we first removed the by-item, and then the by-participant random slopes or, subsequently, random intercepts (following Barr et al., 2013). Effect sizes for comparisons with one fixed effect and random factors were computed following Westfall, Kenny, and Judd (2014). For all between- and within-subjects paired comparisons, Cohen's *d* effect sizes and 95% confidence intervals (CIs) for *d* were used to interpret the magnitude of change, following Larson–Hall and Plonsky (2015).

Table 2 shows the comprehension accuracy by treatment groups. Accuracy on the filler items was high both at pretest and posttest, bearing out that all students were on task.

Table 2 shows that accuracy on subject questions and subject relative clauses was high, while accuracy on object orders was low. Inaccurate answers largely referred to the respective other interpretation of the string, as in subject interpretations of object questions. Answers targeting the animal named in the question (e.g., the horse in [3]) were rare. Therefore, Table 3 gives only accurate responses for majority language and minority language students.

For *wh*-questions and relative clauses, we fitted separate mixed linear logistic regressions for accuracy, with order (subject vs. object), group (PTL vs. control), and test (pretest vs. posttest) as fixed effects including their interactions. Participant and item were included as crossed random factors with random intercepts and uncorrelated random slopes for order and structure. Due to convergence issues, we needed to remove the interaction of order and structure as random slopes. Table 4 gives the results for *wh*-questions and relative clauses.

For both *wh*-questions and relative clauses, significant main effects of order were obtained in the analysis involving both groups, which reflect lower accuracy for object-first orders. For *wh*-questions, the model returned a marginally significant three-way interaction between order, test and group as well as a significant two-way interaction between test and group and a main effect of test. For relative clauses, there were no further significant effects beyond the main effect of order. As seen in the rightmost columns of Table 4, further analyses for *wh*-questions by group show that only the PTL group showed an interaction between order and test, $\beta = 1.21$, $SE = 0.38$, $z = 3.22$, $p = .001$, illustrating that gains from pretest to posttest were specific to object questions in the PTL group ($M_{diff} = 14.8\%$; control group: $M_{diff} = 5.8\%$). In contrast, the control group evinced a main effect of test, which reflects a general moderate increase in performance from pretest to posttest, as is also

TABLE 3
Study 1: Mean Accuracy Scores in % (Standard Deviation in Parentheses) by Test, Construction, and Language Status Group ($n = 114$)

		PTL Group ($n = 62$)				Control Group ($n = 52$)			
		Pretest		Posttest		Pretest		Posttest	
Type	Order	Majority Language ($n = 29$)	Minority Language ($n = 33$)	Majority Language ($n = 29$)	Minority Language ($n = 33$)	Majority Language ($n = 27$)	Minority Language ($n = 25$)	Majority Language ($n = 27$)	Minority Language ($n = 25$)
Fillers		92.1 (8.6)	87.9 (12.2)	95.2 (10.2)	93.6 (9.0)	83.7 (15.0)	87.6 (11.6)	92.6 (9.0)	89.2 (16.3)
Questions	Subject	87.6 (16.4)	86.7 (18.5)	89.0 (21.1)	88.5 (15.8)	85.9 (22.1)	89.2 (16.3)	88.1 (19.4)	91.2 (14.2)
	Object	20.0 (27.3)	13.9 (16.9)	31.7 (37.2)	31.5 (31.6)	16.3 (15.7)	16.8 (23.6)	22.2 (32.5)	22.4 (27.9)
Relative clauses	Subject	92.4 (15.5)	91.5 (15.0)	89.7 (16.6)	81.8 (22.6)	85.2 (21.9)	89.6 (19.3)	85.2 (20.5)	82.4 (27.3)
	Object	24.1 (22.9)	29.1 (22.4)	29.0 (34.9)	24.2 (22.2)	21.5 (21.4)	31.2 (29.5)	26.7 (32.3)	27.2 (31.0)

Note. PTL = pedagogical translanguaging.

seen in the fillers. Critically, these increases were not particular to the target phenomenon of object *wh*-questions. Subsequent analyses that focused on object *wh*-questions returned significant differences between pretest and posttest for the PTL group with a medium effect size, $\beta = 1.06$, $SE = 0.49$, $z = 2.187$, $p = .029$, $d = 0.466$, 95% CI [0.110, 2.012], yet not for the control group, for which the confidence interval of the small effect size crosses zero, $\beta = 0.26$, $SE = 0.52$, $z = 0.488$, $p = .626$, $d = 0.134$, 95% CI [–1.279, 0.769]. Finally, we tested whether there was a difference in gains for object questions between majority language learners and minority language learners in the PTL group. As seen in Table 3, both groups showed gains (majority learners: $M_{diff} = 11.7\%$; minority learners: $M_{diff} = 17.6\%$), although the minority language learners had lower scores at pretest. However, a model with the fixed factors test and language status for object *wh*-questions yielded no interaction between test and language status, $\beta = 0.52$, $SE = 0.67$, $z = 0.768$, $p = .443$.

In sum, the study on *wh*-questions demonstrated that only the PTL group showed learning gains on object *wh*-questions as a result of the instructional treatment. Within the PTL group, both majority language and minority language learners experienced learning gains in comparable measure. There was no transfer of the gains in the comprehension of object questions to syntactically related object relative clauses.

Study 2: Passivization and Wh-*Questions*

Target Phenomenon. In Study 2, we applied PTL to a different phenomenon to explore its scope for word order phenomena that are similar in the FL and the majority language. Like object questions and object relative clauses, passive sentences involve noncanonical mappings between syntactic functions and word order. In a passive sentence, the verb's semantic object is promoted to grammatical subject (4b).

(4) a. *The mother kisses the brother.* (active)
b. *The brother*$_i$ *is kissed* t_i *(by the mother).* (passive)

In English and German, passives are realized similarly by virtue of raising the verb's object to subject and the combination of an auxiliary and the past participle, as seen in the German translations of the English sentences in (5). Due to the basic verb-final word order in German, the optional agentive *by* phrase in long passives occurs in preverbal position.

TABLE 4
Study 1: Glmer Models for Accuracy by Construction and by Group ($n = 114$)

	Both Groups: *Wh*-Questions	Both Groups: Relative Clauses	PTL Group: *Wh*-Questions	Control Group: *Wh*-Questions
Intercept	–2.34***	–1.41***	–2.51***	–2.30***
	(0.35)	(0.28)	(0.38)	(0.34)
Order	4.71***	4.16***	4.86***	4.74***
	(0.56)	(0.55)	(0.59)	(0.53)
Test	0.46*	–0.02	1.45***	0.54**
	(0.25)	(0.23)	(0.27)	(0.26)
Group	–0.01	–0.01		
	(0.40)	(0.33)		
Order* Test	0.15	0.45	1.21***	0.24
	(0.38)	(0.36)	(0.38)	(0.40)
Order* Group	0.03	0.38		
	(0.64)	(0.63)		
Test*Group	0.75**	0.14		
	(0.35)	(0.31)		
Order* Test* Group	0.85+	0.29		
	(0.52)	(0.51)		
Observations	2,280	2,280	1,240	1,040

Note. Cells list estimate and standard error.
$+p < 0.1$. $*p < 0.05$. $**p < 0.01$. $***p < 0.001$.

(5) a. *Die Mutter küsst den Bruder.* (active)
'The mother kisses the brother.'
b. *Der Bruder$_i$ wird (von der Mutter) t_i geküsst.* (passive)
'The brother is (by the mother) kissed.'

Although monolingual children produce passives from an early age, the target comprehension of reversible passives—that is, sentences as in (4) with two animate nouns, in which the word order determines interpretation—is not in place until age 5–6 (e.g., Stromswold et al., 2002). For child L2 learners of English, Marinis (2007) reported that 7–9-year-old L1 Turkish students at UK schools have lower comprehension accuracy than age-matched monolingual students. L1 Turkish-speaking minority language students in Germany also lag behind majority language students at age 7, yet catch up in their comprehension of reversible passives by age 10 (Cristante & Schimke, 2018). In an English-as-a-FL context, passivization is typically introduced only after several years of instruction, and primary school students receive little to no input containing passive sentences.

Instructional Treatments. The instructional treatments were closely modelled after the treatments for questions, using similar contexts and materials. In the first lesson, the teacher used pictures of humans performing actions on inanimate objects, such as a girl hugging a puppet. The teacher then exemplified the contrast between actives and passives by describing the pictures (*the girl hugs the puppet* vs. *the puppet is hugged by the girl*). Students then practiced the descriptions in comprehension and production tasks as in Study 1. In the second lesson, the teacher used the context of the animal birthday party, in which animals hug or kiss each other, so that reversible events needed to be described. As in Study 1, sentences now differed only in word order (*the dog kisses the*

TABLE 5
Study 2: Mean Accuracy Scores and Choices in % (Standard Deviation in Parentheses) by Test, Construction, and Group ($n = 125$)

		Pretest			Posttest		
Type	Order	Accurate	Object/Subject	Named	Accurate	Object/Subject	Named
PTL group ($n = 67$)							
Fillers		84.0 (13.4)			86.4 (11.2)		
Declaratives	Active	87.2 (21.9)	10.1 (19.8)	2.4 (7.4)	86.0 (20.9)	12.2 (19.7)	0.9 (4.2)
	Passive	31.6 (30.4)	64.5 (31.9)	3.6 (8.5)	43.9 (35.7)	54.0 (35.5)	1.5 (5.3)
Questions	Subject	86.9 (20.8)	11.3 (19.1)	0.6 (2.4)	83.3 (21.3)	15.8 (20.5)	0.3 (2.4)
	Object	28.1 (27.9)	67.8 (27.4)	3.6 (9.2)	30.4 (31.0)	66.3 (31.2)	2.1 (6.2)
Control group ($n = 58$)							
Fillers		80.3 (12.8)			89.5 (9.4)		
Declaratives	Active	80.3 (26.6)	12.4 (21.1)	6.6 (12.6)	82.8 (24.4)	13.4 (21.3)	3.4 (10.7)
	Passive	30.7 (29.1)	65.2 (29.4)	1.4 (5.1)	50.0 (31.6)	47.6 (36.9)	1.7 (5.7)
Questions	Subject	77.9 (26.4)	14.1 (21.8)	6.9 (15.7)	79.3 (27.3)	15.2 (22.9)	4.5 (11.9)
	Object	29.7 (26.3)	58.6 (24.8)	7.9 (14.0)	30.0 (31.6)	66.9 (31.9)	2.4 (7.6)

Note. PTL = pedagogical translanguaging; Object/Subject = the respective other interpretation of the string (e.g., subject interpretation of object question or active interpretation of passive sentence); Named = animal named in question.

cat vs. *the cat is kissed by the dog*). Using these pictures, the students practiced the comprehension and production of questions. Subsequently, the teacher in the PTL group wrote down active and passive sentences on the blackboard in English and discussed passives in German. As in Study 1, she asked minority language students to give examples of the translation equivalents of the sentences in other languages spoken by them, noting after each example whether and how the word order differed. Toward the end of the lesson, students filled in a worksheet by writing down active and passive sentences as descriptions of identical events in English, German, and other languages. In the control group, the teacher only wrote down active and passive sentences in English on the blackboard, and the students practiced the use of English sentences further. They completed the same worksheet using only English sentences.

Testing. We used the same test as for Study 1, yet we changed the items for relative clauses into active and passive sentences as in (6 c&d).

(6) a. *Which animal bites the horse?*
b. *Which animal does the horse bite?*
c. *The animal bites the horse.*
d. *The animal is bitten by the horse.*

In all other respects, the test was identical to Study 1.

Testing Procedure. The procedure was identical to the procedure in Study 1.

Participants. One hundred thirty-nine students from Cohort 2 participated in the treatment on passives. Data from both the pretest and the posttest were obtained from 131 students—namely, 69 students in the PTL group (30 minority language learners) and 62 students in the control group (26 minority language learners).

Analysis and Results. We excluded three students, two from the PTL group and one from the control group, since they had English as one of their languages. Three additional participants were removed from analysis because they scored at or below chance on the filler items at both pretest and posttest, suggesting that they did not pay attention to the task. For the remaining 125 participants, we analyzed the data as in Study 1.

Table 5 shows the comprehension accuracy in the respective conditions by groups. As in Study 1, accuracy on the filler items was high for the remaining participants.

All groups displayed high comprehension accuracy on active sentences, while accuracy on passive sentences was below chance at pretest. In the posttest, accuracy on passives increased in both the PTL and the control group, reaching levels around chance. For *wh*-questions, Study 2 replicates the difficulty with object questions found in Study 1, although the overall accuracy on object

questions was higher at pretest among the participants in Study 2. However, there was no change from pretest to posttest in either group for *wh*-questions. Table 6 lists the findings broken down by language status.

For declaratives and *wh*-questions, we fitted separate mixed linear logistic regressions for accuracy, with order (agent/subject-first vs. patient/object-first), group (PTL vs. control), and test (pretest vs. posttest) as fixed effects including their interactions. The random effect structure was as in Study 1. Table 7 presents the models for active or passive sentences and *wh*-questions as well as the by-group models for active or passive sentences.

For active or passive sentences and *wh*-questions, there were significant main effects of order. For *wh*-questions, no further main effects or interactions were found. For active or passive sentences, the model also showed a significant main effect of test as well as an interaction between order and test, indicating that accuracy on passive sentences increased from pretest to posttest. At the same time, there was no significant interaction with group. Further analyses for active or passive sentences by group showed an interaction of order and test for both groups, illustrating that gains from pretest to posttest for passive sentences were obtained for both groups (PTL: $M_{diff} = 12.3\%$; control: $M_{diff} = 19.3\%$). Post-hoc analyses for passive sentences returned significant differences between pretest and posttest for the PTL group, $\beta = 0.79$, $SE = 0.20$, $z = 4.029$, $p < .001$, $d = 0.459$, 95% CI [0.403, 1.168], as well as the control group, $\beta = 1.15$, $SE = 0.21$, $z = 5.422$, $p < .001$, $d = 0.676$, 95% CI [0.731, 1.558], at medium effect sizes. Although the control group showed larger effect sizes, the confidence intervals for neither group crossed zero. Finally, we tested whether there were differential learning gains for passive sentences between majority language learners and minority language learners. A model including both the PTL and the control group with the fixed factors test and language status for passive sentences did not return an interaction between test and language status, $\beta = 0.60$, $SE = 0.40$, $z = 0.151$, $p = .880$, which reflects the results in Table 6—namely, that both majority language and minority language students demonstrated comparable learning gains for passive sentences (majority learners: $M_{diff} = 15.2\%$; minority learners: $M_{diff} = 16.6\%$).

In sum, the study on passive sentences demonstrates that both groups improved in the comprehension of passive sentences following the instruc-

TABLE 6
Study 2: Mean Accuracy Scores in % (Standard Deviation in Parentheses) by Test, Construction, and Language Status Group ($n = 125$)

		PTL Group ($n = 67$)				Control Group ($n = 58$)			
		Pretest		Posttest		Pretest		Posttest	
Type	Order	Majority Language ($n = 40$)	Minority Language ($n = 27$)	Majority Language ($n = 40$)	Minority Language ($n = 27$)	Majority Language ($n = 34$)	Minority Language ($n = 24$)	Majority Language ($n = 34$)	Minority Language ($n = 24$)
Fillers		83.3 (13.5)	85.2 (13.4)	86.3 (9.8)	86.7 (13.3)	79.1 (13.6)	82.1 (11.8)	90.3 (9.7)	88.3 (9.2)
Declaratives	Active	89.5 (19.7)	83.7 (24.8)	89.5 (16.3)	80.7 (25.7)	78.2 (28.9)	83.3 (23.3)	81.2 (26.5)	85.0 (21.5)
	Passive	31.0 (30.4)	32.6 (31.0)	46.0 (38.5)	40.7 (31.6)	32.9 (29.1)	27.5 (29.4)	48.2 (39.7)	52.5 (33.8)
Questions	Subject	90.0 (14.3)	82.2 (27.4)	81.5 (20.9)	85.9 (22.1)	75.3 (28.7)	81.7 (22.8)	74.7 (32.4)	85.8 (16.1)
	Object	25.5 (24.0)	31.9 (32.9)	28.0 (20.3)	34.1 (32.3)	29.4 (23.2)	30.0 (30.6)	28.2 (29.6)	32.5 (34.8)

Note. PTL = pedagogical translanguaging.

TABLE 7
Study 2: Glmer Models for Accuracy by Construction ($n = 125$)

	Both Groups: Active/Passive	Both Groups: *Wh*-Questions	PTL Group: Active/Passive	Control Group: Active/Passive
Intercept	–1.20***	–1.25***	–1.15***	–1.19***
	(0.31)	(0.31)	(0.29)	(0.32)
Order	3.51***	3.05***	4.01***	3.61***
	(0.53)	(0.46)	(0.51)	(0.58)
Test	1.15***	0.08	0.79***	1.15***
	(0.21)	(0.21)	(0.20)	(0.22)
Group	0.05	–0.08		
	(0.36)	(0.33)		
Order* Test	0.93***	0.03	1.00***	0.92***
	(0.34)	(0.31)	(0.34)	(0.34)
Order* Group	0.60	0.76		
	(0.63)	(0.53)		
Test* Group	0.36	0.04		
	(0.29)	(0.29)		
Order* Test* Group	0.07	0.52		
	(0.47)	(0.43)		
Observations	2,500	2,500	1,340	1,160

Note. Cells list estimate and standard error.
$+p < 0.1$. $*p < 0.05$. $**p < 0.01$. $***p < 0.001$.

tional treatment. Neither the PTL nor the control group transferred learning gains in patient-first orders in passives to the comprehension of object *wh*-questions. As in Study 1, majority language and minority language learners experienced comparable learning gains.

DISCUSSION

The two studies reported here assessed effects of plurilingual language teaching on grammatical development in English as a FL among German primary school students. In two short-term teaching interventions embedded in a larger project on plurilingual teaching, we used PTL in instructional treatments for object *wh*-questions (Study 1) and passive sentences (Study 2). The results show larger learning gains for the PTL group for *wh*-questions and comparable gains for the PTL and the control group for passivization.

In both studies, the PTL intervention engendered significant learning gains (RQs 1 and 2), which bears out that the instructional treatments comprising comprehension and production practice were suitable for instruction of the phenomena at hand at the primary school level. At the same time, comprehension accuracy for the target structures remained low at posttest in both experiments, indicating that noncanonical word orders continue to present difficulty among early FL learners at low levels of proficiency who receive limited amounts of FL input (see also Hopp, Steinlen et al., 2019).

In Study 1, both groups had comparable performance at pretest, and the PTL group demonstrated learning gains in the posttest, while the control group did not significantly improve their performance on object *wh*-questions. As the instructional treatments were matched in all respects between the groups except for the inclusion of language comparisons in the PTL

group, the interactions with group and time indicate that cross-linguistic comparisons specifically facilitated grammatical development in the FL. Crucially, the instructional treatments employed cross-linguistic comparisons not as additional elements, but as activities that students completed instead of further practicing the FL structure. Despite such lower time on task in the FL, the PTL group uniquely benefitted from cross-linguistic comparisons in the acquisition of object *wh*-questions that exemplified across-language contrasts between German, on the one hand, and English and other languages, on the other.

These results parallel beneficial effects of explicit instruction on L1 grammatical properties reported in instructed adult L2 acquisition (e.g., McManus & Marsden, 2019b). Unlike in these previous studies, the instructional treatments in our studies did not involve the presentation of explicit rules; instead, they illustrated differences across languages and prompted students to discover cross-linguistic differences and similarities themselves. In these ways, the intervention was appropriate to the ages and proficiency levels of the students and blended with regular classroom and textbook activities. In these contexts, PTL can yield selective advantages over regular FL teaching.

The present findings extend previous research on effects of plurilingual teaching on word formation and vocabulary acquisition (Busse et al., 2020; Leonet et al., 2020) to grammatical development. In conjunction, they highlight that plurilingual teaching methods facilitate FL learning at the primary school level. Although differences in the implementations of PTL and varying degrees of explicitness of language contrasts caution against across-study comparisons, the emerging findings thus suggest that early FL learners benefit from PTL across language domains.

Differential effects of PTL did not surface in Study 2 on passivization, and learning effects did not transfer to other noncanonical grammatical orders in either Study 1 or 2. Hence, effects of PTL were restricted in two ways in the population tested.

First, unlike in word-level studies on plurilingual FL teaching, in which plurilingual advantages arose when the instructional treatments focused on cross-linguistic similarities in affixation or cognate words, the present study found effects of PTL on grammatical development for cross-linguistic differences in the realization of form–function mappings (see also McManus & Marsden, 2019b). Studies 1 and 2 predominantly differed in rationale in that English and German realize the contrast between subject and object *wh*-questions differently, while passives are formed in a largely comparable manner in both languages. A likely explanation for the different findings is that facilitatory effects of PTL predominantly derived from contrasting *wh*-question formation in German versus English and other languages. In contrast, exemplifying similarities in how English and German express passive voice did not engender additive learning over teaching and practicing the formation of passives in English only. Given the similarities between English and German in passivization, we can also not exclude the option that many students in the control group implicitly discovered cross-linguistic parallels and built on them in their learning of English passives. Of course, passivization does differ between English and some minority languages in the classrooms, such as Turkish, in which the passive is formed by verb affixation rather than argument reordering (Özsoy, 2009). However, this did not translate into differential learning effects among the minority language students in the PTL group. We surmise that there may have been too few minority language students who capitalized on these differences, or that many minority students may not have fully acquired passivization in their minority languages for them to be able to draw on their previous linguistic knowledge in the FL (for discussion, Bayram et al., 2019). Future studies should therefore control for the number and type of minority language students and assess the students' knowledge of the respective grammatical phenomena in the minority languages.

Second, the learning gains on object *wh*-questions and on passives did not transfer to object relative clauses in Study 1 or to object *wh*-questions in Study 2. This lack of cross-construction transfer suggests that students did not discover that word order differences in English signal thematic role reversals and could not, therefore, build on this inference in the comprehension of other noncanonical word orders. Hence, the degree to which students generalized grammatical learning was limited. Although the students generalized to novel lexical combinations of verbs and nouns and across different tasks from treatment to testing, their generalizations fell short of more abstract grammatical learning. In both experiments, students appeared to be restricted to pattern-specific or construction learning (see also Lenzing, 2013). It will be interesting to explore if longer instructional treatments or the revision of the phenomena in subsequent lessons can potentially engender greater

abstraction and learning transfer across constructions via inferencing.

In RQ3, we asked whether effects of PTL differed between majority language and minority language students. In neither study did the two groups significantly differ in their performance at either test or between pretest and posttest. These broad parallels in effects of PTL on the two groups could have various reasons. First, parallel development across groups could be due to influence from the majority language German, which was the dominant language for all students. Previous studies in comparable populations of minority language learners have reported that students exhibit grammatical transfer from the dominant language and do not differ from their majority language peers (e.g., Hopp, 2019; Lorenz et al., 2019). In this respect, the inclusion of minority languages may not directly affect grammatical learning differentially among majority and minority language students. Second, many minority language students may not possess sufficient knowledge in their minority languages, in particular for more complex grammatical phenomena, for them to draw on construction-specific linguistic knowledge. Alternatively, their underlying proficiency in the minority language could be too low for facilitative effects to occur more generally (e.g., Cummins, 2017). In fact, anecdotal evidence from the implementation suggests that not all minority language students could supply or verify examples of question or passive formation in their minority languages. Hence, among young minority language speakers who are strongly dominant in the majority language, the direct usefulness of PTL with reference to the majority languages may be restricted to basic differences in, for example, word order. Such differences, however, typically do not present major acquisitional challenges in (early) FL learning—at least in English. Clearly, further research on the scope and the mechanisms of cross-linguistic effects from the minority language is necessary in this population. In any case, the findings suggest that PTL can have additive effects for both minority language and majority language students in mixed primary classrooms.

In sum, the two studies reported here show that plurilingual FL teaching can selectively facilitate grammatical development among primary school students. Despite the reduction in time on target language use, PTL does not seem to entail any learning disadvantages compared to regular FL-only teaching for any learner group.

At the same time, our conclusions are constrained by several limitations. First, we did not test productive knowledge of the phenomena or carry out a delayed posttest, so we could not assess how learning gains would generalize across modalities or time. In addition, the findings are specific to the context and the instructional treatments of this particular study. It is an open question whether positive effects of PTL only surface if, as in this study, the grammatical interventions are embedded in the context of plurilingual teaching over longer periods of time, or whether they also arise if PTL is employed as a targeted teaching method for particular phenomena. In addition, grammatical phenomena are typically not the focus of FL instruction in primary school—for example, passive voice will regularly only be introduced in secondary schooling. It thus remains to be seen to what extent our findings generalize across teaching methods, topics, and domains. Although the majority of students in our studies reported that they enjoyed the plurilingual elements in teaching (Sturm et al., 2021), our observations from the teaching intervention suggest that young FL learners need time, guidance, and habituation to plurilingual teaching for contrastive language activities to proceed smoothly in the classroom (Hopp et al., 2020).

Yet, insofar as the instructional treatments and implementations of PTL in this study represent typical approaches and contexts of plurilingual FL teaching in instructed contexts, the findings show that plurilingual FL teaching can lead to selective advantages over target-language-only instruction in the learning of grammatical contrasts between the FL and the majority language (Study 1). Even when language contrasts are not at issue (Study 2), the lower time on task in the FL inherent to plurilingual teaching did not lead to less learning compared to regular FL-only teaching. Finally, plurilingual teaching benefits all students in mixed classrooms in that both majority language and minority language learners demonstrate selective advantages from across-language comparisons. In conjunction with studies on word-level development, the present findings underscore that PTL is a viable strategy in FL teaching to promote the development of FL skills across language domains and across learner groups. Studies on the efficacy of PTL constitute an important step in validating plurilingual approaches to FL teaching as instructional approaches with a view to meeting the curricular goals in terms of FL proficiency. Beyond fostering advantages in the development of FL skills, plurilingual teaching may have attendant benefits in raising intercultural understanding and learner motivation as well as in engendering awareness and

appreciation of diversity and may thus constitute a key component of an inclusive classroom.

ACKNOWLEDGMENTS

The authors thank the project members, Jenny Jakisch, Teresa Kieseier, and Sarah Sturm, as well as the research assistants, in particular Vivien Schaer, for help with materials development, implementation, data collection, and data coding. The authors are grateful to the project consultant, Carmen Becker, for advice regarding the teaching implementations. The authors thank the editor of *MLJ* and three anonymous reviewers for helpful comments and suggestions. The project MEG-SKoRe II ("Sprachliche und kognitive Ressourcen der Mehrsprachigkeit im Englischerwerb in der Grundschule") was funded by the German Federal Ministry for Education and Research (BMBF) as part of the research cluster "Sprachliche Bildung und Mehrsprachigkeit" (FKZ 01JM1702A/B, 2017–2020).

Open access funding enabled and organized by Projekt DEAL.

NOTES

[1] In the following, we refer to these students as minority language students, even though they are dominant in the majority language, German. The term *minority language* refers to any language that is spoken at home other than and on top of German and that represents a language spoken due to the migration background of the family. It thus does not denote languages that have official minority language status, such as Sorbian, Danish, or others.

[2] In one of the four schools, students also had one English lesson per week in Grades 1 and 2, yet the school followed the same textbook and curriculum in Grades 3 and 4 as the other three schools.

REFERENCES

Arteagoitia, I., & Howard, E. R. (2015). The role of the native language in the literacy development of Latino students in the United States. In J. Cenoz & D. Gorter (Eds.), *Multilingual education* (pp. 61–83). Cambridge: Cambridge University Press.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255–278.

Bates, D., Maechler, M., Bolker, B., Walker, S., Bojesen Christensen, R. H., Singmann, H., … Fox, J. (2019). *Package Lme4: linear mixed-effects models using Eigen and S4* (Version 1.1-21) [Computer software]. https://cran.r-project.org/package=lme4

Bayram, F., Rothman, J., Iverson, M., Kupisch, T., Miller, D., Puig–Mayenco, E., & Westergaard, M. (2019). Differences in use without deficiencies in competence: Passives in the Turkish and German of Turkish heritage speakers in Germany. *International Journal of Bilingual Education and Bilingualism, 22*, 919–939.

Berthele, R., & Udry, I. (2019). Multilingual boost vs. cognitive abilities: Testing two theories of multilingual language learning in a primary school context. *International Journal of Multilingualism*. Advance online publication. https://doi.org/10.1080/14790718.2019.1632315

Bishop, D. V. (2003). *Test for reception of grammar: Test for reception of grammar version TROG-2 stimulus book (1 v.)*. London: Harcourt Assessment.

Bonifacci, P., Canducci, E., Gravagna, G., & Palladino, P. (2017). English as a foreign language in bilingual language-minority children, children with dyslexia and monolingual typical readers. *Dyslexia, 23*, 181–206.

Busse, V., Cenoz, J., Dalmann, N., & Rogge, F. (2020). Addressing linguistic diversity in the language classroom in a resource-oriented way: An intervention study with primary school children. *Language Learning, 70*, 382–419.

Candelier, M., & Kervran, M. (2018). 1997–2017: Twenty years of innovation and research about awakening to languages-Evlang heritage. *International Journal of Bias, Identity and Diversities in Education (IJBIDE), 3*, 10–21.

Cenoz, J. (2003). The additive effect of bilingualism on third language acquisition: A review. *International Journal of Bilingualism, 7*, 71–87.

Cenoz, J. (2013). The influence of bilingualism on third language acquisition: Focus on multilingualism. *Language Teaching, 46*, 71–86.

Cenoz, J. (2019). Translanguaging pedagogies and English as a lingua franca. *Language Teaching, 52*, 71–85.

Cenoz, J., & Gorter, D. (2013). Towards a plurilingual approach in English language teaching: Softening the boundaries between languages. *TESOL Quarterly, 47*, 591–599.

Cenoz, J., & Gorter, D. (2017). Translanguaging as a pedagogical tool in multilingual education. *Language Awareness and Multilingualism. Encyclopedia of Language and Education, 6*, 309–321.

Cenoz, J., & Gorter, D. (Eds.). (2020). Pedagogical translanguaging: Navigating between languages at school and at the university [Special issue]. System, 92.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, the Netherlands: Foris.

Cook, V. J. (2013). Multicompetence. In C. Chapelle (Eds), *The encyclopedia of applied linguistics* (pp. 3,768–3,774). Oxford: Blackwell.

Cristante, V., & Schimke, S. (2018). Verarbeitung von Passiv-und OVS-Sätzen bei Kindern mit Deutsch

als früher L2 [The processing of passive and OVS sentences among children with German as an early L2]. In S. Schimke & H. Hopp. (Eds). *Sprachverarbeitung im Zweitspracherwerb* [Language processing in L2 acquisition], (pp. 169–194). Berlin: DeGruyter.

Cummins, J. (2017). Teaching for transfer in multilingual school contexts. *Bilingual and Multilingual Education*, *1*, 103–116.

DeVincenzi, M. (1991). *Syntactic parsing strategies in Italian*. Dordrecht, the Netherlands: Kluwer Academic Publishers.

Dunn, L. M., Dunn, D. M., & Styles, B. (2009). *British picture vocabulary scale: BPVS-III*. (3rd ed.). London: GL Assessment.

Elsner, D. (2007). *Hörverstehen im Englischunterricht der Grundschule: Ein Leistungsvergleich zwischen Kindern mit Deutsch als Muttersprache und Deutsch als Zweitsprache* [Listening comprehension in English lessons in primary school: A comparison of achievement between children having German as the first and the second language]. Bern, Switzerland: Peter Lang.

Fäcke, C., & Meißner, F. J. (Eds.). (2019). *Handbuch Mehrsprachigkeits-und Mehrkulturalitätsdidaktik*. [Handbook multilingualsm- and multiculturalism for teaching]. Tübingen, Germany: Narr Francke Attempto Verlag.

Fleckenstein, J., Möller, J., & Baumert, J. (2018). Mehrsprachigkeit als Ressource. [Multilingualism as a resource]. *Zeitschrift für Erziehungswissenschaft*, *21*, 97–120.

Friedmann, N., & Novogrodsky, R. (2004). The acquisition of relative clause comprehension in Hebrew: A study of SLI and normal development. *Journal of Child Language*, *31*, 661–681.

García, O., & Li Wei. (2014). Language, bilingualism and education. In O. García & Li Wei (Eds.), *Translanguaging: Language, bilingualism and education* (pp. 46–62). London: Palgrave Macmillan.

Gerngross, G., Puchta, H., & Becker, C. (2013). *Playway 4*. Rum/Innsbruck, Stuttgart: Helbling; Klett.

Göbel, K., & Vieluf, S. (2017). Specific effects of language transfer promoting teaching and insights into the implementation in EFL-teaching. *Orbis Scholae*, *11*, 103–120.

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, *36*, 3–15.

Haenni Hoti, A. U., Heinzmann, S., Müller, M., Oliveira, M., Wicki, W., & Werlen, E. (2011). Introducing a second foreign language in Swiss primary schools: The effect of L2 listening and reading skills on L3 acquisition. *International Journal of Multilingualism*, *8*, 98–116.

Hesse, H.–G., Göbel, K., & Hartig, J. (2008). Sprachliche Kompetenzen von mehrsprachigen Jugendlichen und Jugendlichen nicht-deutscher Erstsprache [Language competencies of multilingual adolescents and adolescents with non-German first languages]. In DESI-Konsortium (Ed.). *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* [Instruction and achievement in German and English: Results from the DESI study] (pp. 208–230). Weinheim, Germany: Beltz, Empirische Bildungsforschung.

Hirosh, Z., & Degani, T. (2018). Direct and indirect effects of multilingualism on novel language learning: An integrative review. *Psychonomic Bulletin & Review*, *25*, 892–916.

Hopp, H. (2019). Cross-linguistic influence in the child third language acquisition of grammar: Sentence comprehension and production among Turkish–German and German learners of English. *International Journal of Bilingualism*, *23*, 567–583.

Hopp, H., Jakisch, J., Sturm, S., Becker, C., & Thoma, D. (2020). Integrating multilingualism into the early foreign language classroom: Empirical and teaching perspectives. *International Multilingual Research Journal*, *14*, 146–162.

Hopp, H., Steinlen, A., Schelletter, C., & Piske, T. (2019). Syntactic development in early foreign language learning: Effects of L1 transfer, input, and individual factors. *Applied Psycholinguistics*, *40*, 1,241–1,267.

Hopp, H., Vogelbacher, M., Kieseier, T., & Thoma, D. (2019). Bilingual advantages in early foreign language learning: Effects of the minority and the majority language. *Learning and Instruction*, *61*, 99–110.

Hufeisen, B., & Neuner, G. (2003). *Mehrsprachigkeitskonzept-Tertiärsprachenlernen-Deutsch nach Englisch*. [Concept of plurilingualism—Third language learning—German after English]. Strasbourg, France: Council of Europe Publishing.

Kasprowicz, R. E., Marsden, E., & Sephton, N. (2019). Investigating distribution of practice effects for the learning of foreign language verb morphology in the young learner classroom. *Modern Language Journal*, *103*, 580–606.

Krumm, H.–J., & Reich, H. (2013). *Multilingualism curriculum: Perceiving and managing linguistic diversity in education*. Münster, Germany: Waxmann.

Larson–Hall, J., & Plonsky, L. (2015). Reporting and interpreting quantitative research findings: What gets reported and recommendations for the field. *Language Learning*, *65*, 127–159.

Lenzing, A. (2013). *The development of the grammatical system in early second language acquisition: The multiple constraints hypothesis* (Vol. 3). Amsterdam: John Benjamins.

Leonet, O., Cenoz, J., & Gorter, D. (2020). Developing morphological awareness across languages: Translanguaging pedagogies in third language acquisition. *Language Awareness*, *29*, 41–59.

Lorenz, E., Bonnie, R. J., Feindt, K., Rahbari, S., & Siemund, P. (2019). Cross-linguistic influence in unbalanced bilingual heritage speakers on subsequent language acquisition: Evidence from pronominal object placement in ditransitive clauses. *International Journal of Bilingualism*, *23*, 1,410–1,430.

Lyster, R., Quiroga, J., & Ballinger, S. (2013). The effects of biliteracy instruction on morphological

awareness. *Journal of Immersion and Content-Based Language Education*, *1*, 169–197.

Maluch, J. T., & Kempert, S. (2017). Bilingual profiles and third language learning: The effects of the manner of learning, sequence of bilingual acquisition, and language use practices. *International Journal of Bilingual Education and Bilingualism*, *22*, 870–882.

Maluch, J. T., Kempert, S., Neumann, M., & Stanat, P. (2015). The effect of speaking a minority language at home on foreign language learning. *Learning and Instruction*, *36*, 76–85.

Maluch, J. T., Neumann, M., & Kempert, S. (2016). Bilingualism as a resource for foreign language learning of language minority students? Empirical evidence from a longitudinal study during primary and secondary school in Germany. *Learning and Individual Differences*, *51*, 111–118.

Marinis, T. (2007). On-line processing of passives in L1 and L2 children. In *Proceedings of the 2nd conference on Generative Approaches to Language Acquisition North America (GALANA)* (pp. 265–276). Somerville, MA: Cascadilla Proceedings Project.

Martinez, H. (2015). Mehrsprachigkeitsdidaktik: Aufgaben, Potenziale und Herausforderungen. [Plurilingual teaching: Aims, potentials and challenges]. *FLuL–Fremdsprachen Lehren und Lernen*, *44*, 7–19.

McManus, K., & Marsden, E. (2017). L1 explicit instruction can improve L2 online and offline performance. *Studies in Second Language Acquisition*, *39*, 459–492.

McManus, K., & Marsden, E. (2019a). Signatures of automaticity during practice: Explicit instruction about L1 processing routines can improve L2 grammatical processing. *Applied Psycholinguistics*, *40*, 205–234.

McManus, K., & Marsden, E. (2019b). Using explicit instruction about L1 to reduce crosslinguistic effects in L2 grammar learning: Evidence from oral production in L2 French. *Modern Language Journal*, *103*, 459–480.

Montrul, S. (2016). *The acquisition of heritage languages*. Cambridge: Cambridge University Press.

Niedersächsisches Kultusministerium. (2018). *Kerncurriculum für die Grundschule Schuljahrgänge 3–4: Englisch*. Hannover. Accessed 7 May 2021 at http://db2.nibis.de/1db/cuvo/datei/2018-06-25_kc_englisch_gs_druckfassung.pdf

Özsoy, S. (2009). Argument structure, animacy, syntax and semantics of passivization in Turkish: A corpus-based approach. In Y. Kawaguchi, M. Minegishi, & J. Durand (Eds.), *Corpus analysis and variation in linguistics* (pp. 259–279). Amsterdam: John Benjamins.

Petermann, U., & Petermann, F. (2008). HAWIK-IV. *Kindheit und Entwicklung*, *17*, 71–75.

R Development Core Team. (2020). R (Version 3.6.3) [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. https://www.r-project.org/

Rankin, T. (2014). Variational learning in L2: The transfer of L1 syntax and parsing strategies in the interpretation of *wh*-questions by L1 German learners of L2 English. *Linguistic Approaches to Bilingualism*, *4*, 432–461.

Rinker, T., & Gagarina, N. (2014). *Cross-linguistic lexical task: German version (CLT-DE)*. Unpublished instrument. Accessed 25 April 2021 at https://multilada.pl/en/projects/clt/

Roesch, A. D., & Chondrogianni, V. (2016). "Which mouse kissed the frog?" Effects of age of onset, length of exposure, and knowledge of case marking on the comprehension of *wh*-questions in German-speaking simultaneous and early sequential bilingual children. *Journal of Child Language*, *43*, 635–661.

Schoonen, R., Gelderen, A. V., Glopper, K. D., Hulstijn, J., Simis, A., Snellings, P., & Stevenson, M. (2003). First language and second language writing: The role of linguistic knowledge, speed of processing, and metacognitive knowledge. *Language Learning*, *53*, 165–202.

Stromswold, K., Eisenband, J., Norland, E., & Ratzan, J. (2002, March). Tracking the acquisition and processing of English passives: Using acoustic cues to disambiguate actives and passives. Paper presented at the CUNY conference on sentence processing, New York. March 21.23, 2002. Accessed 25 April 2021 at https://ruccs.rutgers.edu/images/personal-karin-stromswold/publications/Kstrompsv.pdf

Sturm, S., Jakisch, J., Kieseier, T., Thoma, D., & Hopp, H. (2021). Pedagogical translanguaging in early foreign-language teaching and its effects on minority-language and majority-language students. In S. Dietrich–Grappin, B. Hufeisen, & Li Wei (Eds.), *Vers la compétence plurilingue: Translanguaging im schulischen Tertiärsprachenunterricht*. Frankfurt: Peter Lang.

Turnbull, B. (2018). Reframing foreign language learning as bilingual education: Epistemological changes towards the emergent bilingual. *International Journal of Bilingual Education and Bilingualism*, *21*, 1,041–1,048.

Weiß, R. H. (2006). *CFT 20-R. Grundintelligenztest Skala 2: Manual*. Göttingen: Hogrefe.

Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General*, *143*, 2,020–2,045.

Wilden, E., & Porsch, R. (2016). Learning EFL from year 1 or year 3? A comparative study on children's EFL listening and reading comprehension at the end of primary education. In M. Nikolov (Ed.), *Assessing young learners of English: Global and local perspectives* (pp. 191–212). Cham, Switzerland: Springer.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.